

Gender and the evaluation of research

Article

Accepted Version

Brooks, C., Fenton, E. M. and Walker, J. T. (2014) Gender and the evaluation of research. *Research Policy*, 43 (6). pp. 990-1001. ISSN 0048-7333 doi:
<https://doi.org/10.1016/j.respol.2013.12.005> Available at
<https://centaur.reading.ac.uk/36810/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1016/j.respol.2013.12.005>

To link to this article DOI: <http://dx.doi.org/10.1016/j.respol.2013.12.005>

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

NOTICE: this is the author's version of a work that was accepted for publication in research in the Journal of Empirical Finance. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Research Policy, Volume 43, Issue 6, July 2014, Pages 990–1001.

DOI: doi:10.1016/j.respol.2013.12.005

Gender and the Evaluation of Research

Chris Brooks*, Evelyn M. Fenton** and James T. Walker***

* ICMA Centre, Henley Business School, University of Reading
(C.Brooks@reading.ac.uk)

**Leadership, Organisations and Behaviour, Centre for Social and Organisational Studies,
Henley Business School, University of Reading
(e.m.fenton@henley.reading.ac.uk)

***International Business and Strategy, Centre for International Business History,
Henley Business School, University of Reading
(Corresponding author - j.t.walker@henley.reading.ac.uk)

Word count: 9,331 (7,252 body text)

Abstract: This study examines if and how gender relates to research evaluation via panel assessment and journal ratings lists. Using data from UK business schools we find no evidence that the proportion of women in a submission for panel assessment affected the score received by the submitting institution. However, we do find that women on average receive lower scores according to some journal ratings lists. There are important differences in the rated quality of journals that men and women publish in across the sub-disciplines with men publishing significantly more research in the highest rated accountancy, information management and strategy journals. In addition, women who are able to utilise networks to co-author with individuals outside their institution are able to publish in higher-rated journals, although the same is not true for men; women who are attributed with “individual staff circumstances” (e.g. maternity leave or part-time working) have lower scores according to journal ratings lists.

Keywords: Research quality, research assessment exercise, research excellence framework, research evaluation, gender.

Acknowledgements: The authors thank Tom Briddle and Will Harvey for their research assistance and three anonymous referees of this journal as well as Adrian Bell, Alan Rugman and Anna Vignoles for very useful comments. The first author was a member of the Accounting and Finance Sub-Panel for RAE2008 and is a member of the Business and Management Sub-Panel for REF2014. We also thank our schools, the ICMA Centre and International Business and Strategy, for funding this work.

1. Introduction

Evaluating the quality of academic research and its policy outcomes is an important task. Academic research is under pressure to become more relevant to society and to pay its way (Nightingale and Scott, 2007), while the knowledge-based economy is seen as key to national competitiveness (Department for Business Innovation and Skills, 2012). Providing and developing an internationally competitive higher education sector, with high-quality teachers and researchers, is central to building such an economy and is a core objective of national and regional entities (e.g. European Commission, 2005).

Assessing new knowledge derived via research is controversial, and how that evaluation occurs may have important impacts. Gender equality is also a pressing policy issue in higher education (Jacobs and Winslow, 2004; London et al., 2012; Long et al., 1993; Mayer and Tikka, 2008; Miller et al., 2005; Rama et al., 1997; Van Den Brink et al., 2006; Wolfinger et al., 2009), especially since legislation requires the public sector to promote equality rather than simply prevent discrimination.¹ There has been a shift internationally toward more ratings-based methods to measure the quality and impact of research, alongside other methods such as research assessment exercises (Butler, 2007; Coupé et al., 2010; Donovan, 2007; Moed, 2007).² Evaluation apportions not only income³ but reputational impact on both the school and the individuals within it. The use of “A-journal” listings forms a critical element of US and Canadian tenure decisions, although these listings are not publicly provided and differ across institutions.⁴ In Australia, institutions now follow a combination of indicators and expert review by committees comprising experienced, internationally recognised experts, having moved away from

¹For example, the Equality Bill (2007).

²Some care is necessary since the term “impact” as it is used within the bibliometrics literature relates purely to numbers of citations. Since citations can be positive or negative (e.g., papers are cited as exemplars of poor research or to correct their errors), it is important to distinguish impact from “quality” as the term is used within the RAE/REF in the UK, where it is taken to mean the originality, significance and rigour of a piece of work.

³ According to the Higher Education Funding Council for England, around £1.5bn was allocated in “QR” funding, which depends on research quality, to English universities alone in 2011–12.

⁴ Most internationally recognised US and Canadian institutions require six “A-list” publications over six years to obtain tenure. The specific outlets that make up the A-list differ between institutions but rankings form the basis of these.

explicitly incorporating journal rankings (Donovan, 2007).⁵ New Zealand has the Performance Based Research Fund (PBRF), developed to encourage excellent research at institutions of higher education. A total of 60% of the research funding available through the PBRF is allocated according to a peer assessment of individual research performance, which rates individuals' portfolios as having international, national or local standing.⁶ Italy is developing a research exercise similar to the UK's.⁷ There are also a host of alternative rating lists relating to business and management, including the Erasmus Research Institute of Management list, the ESSEC Research Centre Ranking of Journals, and subject-specific ratings such as that produced by the Tinbergen Institute and Internal Kiel Institute Journal Ranking.⁸

In the UK, the Research Assessment Exercise (RAE, now renamed the Research Excellence Framework (REF) for 2014) provides a comprehensive evaluation of the quality of research.⁹ The RAE 2008 was the most recent of four evaluations conducted by academic institutions in the UK.¹⁰ The objective was to produce quality profiles and sub-profiles for each submitting unit (Broadbent, 2010). The assessments were performed by panels of experts who produced quality profiles in three areas: outputs, environment and esteem.

⁵ See also http://arc.gov.au/era/era_2012/era_2012.htm for the specific arrangements of the most recent exercise conducted in 2012.

⁶ Detail concerning the PBRF is found at <http://www.tec.govt.nz/Funding/Fund-finder/Performance-Based-Research-Fund-PBRF/>.

⁷ <http://www.anvur.org/?q=en/content/procedura-di-valutazione>.

⁸ The Erasmus Research Institute of Management list is found at <http://www.erim.eur.nl/ERIM/About/EJL>; <http://www.tinbergen.nl/research-institute/journal-list.php>; the ESSEC Research Centre Ranking of Journals is found at http://www.essec.edu/fileadmin/user_upload/Rubrique_Professeurs_et_recherche/Recherche/revues-management-classification.pdf; that of the Tinbergen Institute at (<http://www.tinbergen.nl/research-institute/journal-list.php>); and the Internal Kiel Institute Journal Rankings at (<http://www.ifw-kiel.de/forschung/internal-journal-ranking>).

⁹ The RAE was an evaluation of the quality of research produced by UK universities run jointly by the Higher Education Funding Council for England (HEFCE), the Scottish Funding Council (SFC), the Higher Education Funding Council for Wales (HEFCW), and the Department for Employment and Learning, Northern Ireland (DEL). Any higher education institute in the UK that is eligible to receive research funding from one of these bodies was eligible to participate in the RAE, and the evaluation was done separately by subject area. The results of the exercises have been used to determine the amount of QR funding allocated to universities for their research. The exercise was conducted in 1986, 1989, 1992, 1996, 2001 and 2008. The RAE has now been replaced by the REF, which has similar objectives and will operate much like the RAE: see www.rae.ac.uk.

¹⁰ Individual higher education institutions were free to select which subject area sub-panels to submit their work to for the RAE, and it might be that two or more schools from a given institution were combined together in a single submission. Therefore, in this paper, we use the terms "school", "department", "institution", "submitting unit" and "university" interchangeably since almost invariably each institution would make a maximum of one submission to the Business and Management Sub-Panel.

Here we explore whether and how the gender of an author affects evaluations of research quality in business and management in the UK. Business and management is a significant cognate area regarding research evaluations, as the magnitude of business schools relative to other university departments means they are subject to much of the pressure for improved scores by academic institutions (Piercy, 2000). The UK is a particularly suitable setting for this enquiry, as it has a comparatively homogeneous higher education system, and a long history of research assessment (Collini, 2008). Moreover, the UK has witnessed repeated concerns about the use of journal ratings lists (Morris et al., 2011; Northcott and Linacre, 2010; Oswald, 2007; Piercy, 2000). Following considerable debate, the REF administrators have stated that in REF2014, no sub-panel will exclusively employ metrics to evaluate research quality,^{11, 12} in part due to suggestions that they may disadvantage women (HEFCE, 2011).¹³

As is the case internationally, the use of journal ratings lists has become increasingly popular in the UK. It is seen as a means to objectify research assessment and avoid or compensate for any biases in peer review (Taylor, 2011a). Yet journal list-based evaluation has been criticised as inappropriate for this role, being inadequate as a measure of journal quality (Easton and Easton, 2003; Moed, 2007) where the indicator becomes a target leading to gamesmanship by the academic (Macdonald and Kam, 2007a, b) and leaving lower-rated journals struggling with diminished quality and quantity of submissions (Northcott and Linacre, 2010). Yet the proliferation of journal rating lists indicates increasingly wide usage across disciplines (both explicitly and implicitly) for a variety of quality assessment purposes, such as resourcing,

¹¹ REF2014: Assessment Framework and Guidance on Submissions, document no. 2011–12, p. 4, states that “while these experts will draw upon appropriate quantitative indicators to support their professional judgement, expert review remains paramount.” Even more explicitly, the Panel Criteria and Working Methods document for Main Panel C, p.64, states that, “No sub-panel will use journal impact factors or any hierarchy of journals in their assessment of outputs.” In addition, the Business and Management Sub-Panel elected to neither receive nor make use of citation data for individual outputs (p.66). See www.ref.ac.uk/subguide/citationdata/contextualdata.

¹² We recognise the distinction between “bibliometrics,” which usually refer to an assessment of the citation scores of a specific output, and journal impact factors. However, we employ data on journal ratings rather than those of the individual studies published in them, following the procedure that many institutions appear to be adopting in preparation for the REF.

¹³ The study utilises information from 22 institutions across 35 RAE units of assessment drawing upon the Web of Science and Scopus databases. The study finds there is a gender difference comparing men’s and women’s scores. It does not consider the mechanisms by which women are disadvantaged.

recruitment, merit raises and promotion (Agrawal et al., 2011; Giles and Garand, 2007; Reinstein and Calderon, 2006; Voss, 2010). The Association of Business Schools (ABS) and the *Financial Times* journal lists are the most actively employed UK listings. The former has achieved some currency in other countries, where citation-determined impact factors also commonly appear in workload models, and feature in the discussions of interview panels and promotion committees (Beattie and Goodacre, 2012). Studies have demonstrated that the journal ratings of a department's publications are the strongest predictor of the results obtained in the 2008 UK's RAE, although journal ratings were not formally used in the evaluation (Kelly et al., 2009; Taylor, 2011b). University managers appear to be making increasing use of such journal ratings to prepare for the forthcoming assessment (REF2014).

Examining how gender relates to research assessment is complicated by the number of mechanisms through which women may be affected (Dwyer, 1994; Ginther and Kahn, 2004; Maranto and Griffin, 2011; Probert, 2005; Ward, 2001a). However, it is not clear if the use of different measures affects women, or how they may reinforce and/or interact with each other. While there have been a number of qualitative studies (Aksnes et al., 2011; Haynes and Fearfull, 2008; Knights and Richards, 2003) on gender effects, there has been less work examining how gender directly affects research exercises involving a peer review process or indirectly affects research evaluations via citation or journal ratings lists.

This paper examines both panel processes and journal ratings lists, which measure the quality of research in business and management in the UK. It makes two contributions. First, it examines whether the proportion of women in institutions submitted to the Business and Management Unit of Assessment in RAE 2008 impacts on institutional performance. Second, we examine whether journal ratings lists could indirectly affect the measured performance of women. We also examine whether indirect effects from the extent of co-authorship and individual circumstances, such as maternity leave, affect women's performance, and how much journal

ratings lists differ between and within subject areas in business and management given the uneven distribution of women across sub-disciplines.

The remainder of this paper is organised as follows. Section 2 develops a conceptual framework that identifies possible mechanisms through which differences between measured research quality across genders may occur within a peer review panel and/or via journal ratings lists. Section 3 discusses the methods that we use to analyse various measures of quality and Section 4 provides our results. Section 5 discusses the implications of the key findings relating to gender and Section 6 offers some concluding comments including the relevance of our results for other disciplines and implications for research policy.

2. Conceptual framework and hypotheses

We identify four potential mechanisms at both the institutional and individual level through which women may be affected by the nature of research evaluation.

2.1 Institutional-level mechanisms

The evaluation of research quality in the UK follows a panel framework supplemented by journal ratings lists in some disciplines. Women could therefore be affected directly via the evaluations of the panel members. While we are unaware of work suggesting gender effects, recent work has argued that there is panel bias at the institutional level with a “halo” effect at work that privileges universities with reputational capital (Colyvas, 2012). Peer review is conducted via panel members reading and evaluating authors’ research. Reviewers have access to the name, affiliation, and gender of each author, allowing the possibility of differential treatments of men’s and women’s work. Men outnumber women substantially in business and management with only 25% of participants in the 2008 research exercise being women. This skewed representation leads to predominantly male panellists, and although this does not in itself imply any sort of bias, it remains to be tested.

Hypothesis 1: Institutions with a higher proportion of women will have lower ratings than those that do not.

Does the proportion of women on a faculty affect the outcome of the RAE? The gender of each author of a piece of research is known, and this may affect the assessment, or it may be that women concentrate on sub-fields that produce higher or lower scores than other sub-fields, or that women produce work of a systematically different standard from men's.

Hypothesis 2: Institutions with a higher proportion of women who co-author will have higher ratings than those that do not.

While identifying potential gender effects on RAE outcomes is an important first step, unpacking the mechanism(s) through which these effects may occur is of key policy relevance. One key mechanism may be what earlier work has suggested to be the existence of “old boy” networks that affect the joint decision to co-author and to publish (Jordan et al., 2008). Research in this area is contradictory: in the US, women in sociology (Leahey et al., 2008) and more generally (Bozeman and Gaughan; 2011) collaborate more than men; but McDowell et al.(2005) find that networks have effects that differ by gender. Women may be less able to engage in the type of gamesmanship required to develop broad external networks of co-authorship and citation (Pezzoni et al., 2012).

Hypothesis 3: Institutions with a higher proportion of women who have “individual staff circumstances” such as maternity leave, part-time working, or were classified as early career researchers, will have lower ratings than those who do not.^{14,15}

¹⁴Given the data source that we employ, and data protection issues, the precise nature of individual staff circumstances are not made public, so we are not able to disentangle the various categories of circumstances (e.g. maternity versus only just having joined the profession).

There are several reasons why women may be less productive than their male counterparts, resulting in fewer outputs to select from to enter for research assessments. Inequalities in career progress have been explained by the greater mobility of men (Ward, 2001a, b), perceptions of exclusion (Maranto and Griffin, 2011), and the lack of senior female role models (Sealy and Singh, 2010). Additionally, women are impacted by what the economics literature terms “the family gap” or “the child penalty” (Budig and England, 2001; Waldfogel, 1998a, b; Ward, 2001a, b) and suffer from not only a productivity decline that subsequently rises less quickly for women than men following children (Leahey et al., 2008) but also reduced visibility (Hunter and Leahey, 2010). This last may result in women having different bargaining power, exacerbated by stronger geographic ties due to the expense of moving children. The family gap may exist for several reasons: maternity and parenthood provide real productivity shocks through sleep deprivation and an increased propensity for illness in the household; they may delay the accumulation of human capital (Pezzoni et al., 2012), and reduce the time and mental space required for concentration; and mothers may invest differently in household production resulting, for example, in career interruptions. Men by contrast are said to gain from parenthood (Browning, 1992; Millimet, 2000; Simonsen and Skipper, 2008). Furthermore, women are more likely to work part-time, which may have an effect on the quantity and the quality of output. Women are also more likely to be responsible for elderly parents or recalcitrant teenagers (Probert, 2005).

Hypothesis 4: Institutions with a higher proportion of staff in sub-disciplines highly represented by women will have lower ratings than those that do not.

Academics can be wedded to their sub-discipline because of sunk costs in terms of accumulated intellectual capital and reputational assets. Some authors have argued that aspects of previous RAEs exhibit a systematic bias against some disciplines. Stewart (2005) records the detrimental effects of the concepts used to judge outputs in the area of human resource development, while

¹⁵ An early career researcher was defined as someone who had entered the profession within the RAE assessment period between January 2001 and October 2007.

Rafols et al. (2012) find journal ratings detrimental to the overall scores for innovation studies. Table 1 shows that women tend to be clustered in particular sub-disciplines such as human resource management and marketing.

2.2 Individual-level mechanisms

While gender effects may exist at the panel level, it is also possible that commonly employed journal ratings lists are gendered. Studies looking at social science disciplines have suggested that the use of impact factor measures is biased against women (e.g. Davenport and Snyder, 1995; McElhinny et al., 2003; Ward et al., 1992).

Although the rules underlying research exercises are transparent, journal ratings lists are usually not produced in an entirely transparent way. Those producing the lists are typically neither elected nor accountable, either for the journal ratings themselves or for the decisions that follow. The motivation behind their development is also ambiguous. For example, it is not clear whether the aim of the most actively employed UK listing, the ABS list, was to capture on aggregate, or even supplant the research assessment process. The list was first produced around the time when institutions were preparing their submissions for the RAE2008, and the journal ratings are provided on a similar 4-point scale.¹⁶

Journal ratings lists may have substantial impacts on academic journals and disciplines. It has been implied that journal ratings lists distort perceptions of quality across sub-fields (Easton and Easton, 2003; Findlay and Sparks, 2010; Morris et al., 2011; Stewart, 2005). As we demonstrate below, given that the participation of women is unevenly distributed across sub-disciplines, such distortions may have consequences. The percentage of journals rated as 3 and 4 by the ABS list varies considerably across these sub-fields for reasons that are not entirely obvious and do not appear to relate specifically to the relative qualities of the sub-field's work.

¹⁶ It appears that the implications of each quality rating according to the ABS list (see Morris et al., 2009, Table I) suggest the use of somewhat less stringent criteria than those used by the RAE sub-panels when forming their judgements, and this is reflected in the higher proportion of work that would have been rated at 4* by the ABS list than by the RAE 2008 sub-panel in Business and Management Studies.

Consequently, when such ratings are used uncritically, they could lead to misallocations of resources.¹⁷ Such issues are likely to become more serious as the use of lists appears to be increasingly prevalent and decisions made as a result of list usage become more embedded within institutional structures.

Hypothesis 1.1: Women's outputs will be published in journals with lower ratings than those of men.

Arguably a similar set of factors to hypothesis 1 may also indirectly affect women, leading us to posit a further set of hypotheses at the individual level.

Hypothesis 2.1: Women who co-author will have higher journal ratings than those who do not.

The potential for network effects identified as underpinning hypothesis 2 will have a similar impact at the individual level of analysis in that the ability to network and hence to co-author externally creates intellectual synergies leading to better quality outputs for women than if writing on their own. Women whose mobility is hampered by child care responsibilities or who do not find suitable co-authors may have fewer opportunities for networking (Jordan et al., 2008).

Hypothesis 3.1: Women who have fewer than four outputs due to "individual staff circumstances" will have lower quality journal ratings for their remaining outputs.

We have discussed how women's roles as mothers and carers may impact upon their outputs for the institution as a whole. The output of women who are classified as having individual staff circumstances may not be just fewer in number, but also of lower quality. Women with time out may have spent less time at conferences and seminars; they do not benefit from interactions with

¹⁷For example, the letter written by Christine Helliard argues that several key journals in accounting have been misclassified, and worse, reclassified to a far less appropriate sub-field than they were in a previous iteration of the ABS guide. The ratings of these journals were significantly reduced with potentially disastrous consequences for the numbers of papers submitted to them. The letter can be found at: <http://bafa.group.shef.ac.uk/>; see also Hoepner and Unerman (2012).

and inputs from colleagues, and miss out on network contacts. They may also suffer from an inability to spend concentrated periods of time engaged in the thinking required to formulate a high-quality journal article.

Hypothesis 4.1: Women who publish in sub-disciplines with a high representation of women will have lower quality journal ratings.

Gender differences in citation rates appear to be sub-discipline specific, so identifying whether a difference exists within a discipline is an important factor for making fair and equitable decisions regarding the evaluation and promotion of female versus male researchers (Ward et al., 1992).¹⁸ This is of particular relevance to a multi-disciplinary area such as business and management. The participation rates of women vary from between 14% in finance to 36% in human resource management. Where journal ratings lists have a predominance of particular subject areas, or alter their coverage to reduce or incorporate outlets in certain sub-disciplines, these may harm or benefit individuals who specialise.

3. Methods

We developed a set of variables at the institutional and individual levels in order to examine the direct effect of gender on the UK Research Assessment Exercise for Business and Management. The data that we use in this study come from the RAE website.¹⁹

3.1 Measures of “quality”

We calculate the “grade point average” (GPA) for research output in the current standard fashion of multiplying each quality rating (on a 0 to 4 scale) by the percentage of work graded to be in that category. We do not present summary statistics for the scores, since such analyses are already

¹⁸There was also discussion of the perceived unequal opportunities in the press (Times Higher Education, 2009).

¹⁹<http://www.rae.ac.uk/submissions/download.aspx?option=uoa>

available in Kelly et al. (2009). Although 90 universities made returns to the Business and Management unit of assessment, Buckinghamshire New University entered only two category A faculties, and therefore, in line with the RAE2008 rules, their overall research quality profile was made public but the component profiles for outputs, environment and esteem were not. Hence our sample for most of the analysis focuses on the GPA scores for the output indicator for the remaining 89 institutions that constitute our aggregate-level sample.

RAE panel results are only made available at the institution level. Therefore, School Directors of Research who want to determine the quality of the research conducted by specific individuals or groups of researchers within their institution must rely on some other measure. So, as well as the grade point averages for the output we include the average rating according to the four most widely invoked measures in the discipline: the ABS list, citation based “impact factors” from the ISI database, and the *Financial Times*(FT) 40 and 45 lists.

In the UK, the most commonly used journal rating list in business and management is that produced by the Association of Business Schools (ABS) described by Morris et al. (2009, p. 1446): as “...a hybrid list...developed through an iterative process. ...[It] originated from a list of all journals from which three or more articles were submitted to the business and management panel at RAE2001. Other journals were then added through comparison with lists from six UK business schools: Aston, Cranfield, Durham, Imperial, Kent and Warwick... The next stage was to compute a citation impact factor on a four point scale... by taking the mean citation impact factor for the last three years.” However, many journals do not have published impact factors and therefore ABS rated other journals based on an examination of the journal’s website, consideration of its editorial policies, and the quality of articles in three recent editions. The journal ratings produced using this procedure were then moderated according to the opinions of researchers in each of 22 sub-discipline areas.

The FT produced a list of 40 journals spanning various sub-disciplines in business and related areas. Each publication in those journals increases the research score component of

various ratings produced by the FT, such as their table of MBA programmes. The journal list was updated in 2010 and, while the overall number of journals increased to 45, some journals that were included in the original list were replaced (e.g. *International Journal of Human Resource Management*).

The use of these measures can influence hiring, pay, sabbatical time and workload models, so the shift of journals incorporated in the FT45, compared with the FT40 list, provides a means to examine how a re-classification of journals can affect subject areas. We exclude all non-journal RAE outputs since they are not included in any of the journal or impact factor ratings that we utilise.²⁰

For the institution-level analysis we measure dispersion in the quality of entries according to the ratings lists, which has a separate value for each submitted member of staff (which we then average within an institution) and we are also interested in examining whether more dispersed individual entries reduce the average measured quality. An individual with only one high-quality output, as judged by the ABS and impact factor measures, may fare worse than those who were more consistent. We adopt a simple range statistic for the ABS scores and measure the interquartile range for the impact factor, because the ABS score, being between one and four, exhibits less variation than the impact factor measure, which is heavily skewed, and we want to capture these differences.

3.2 Individual circumstances and co-authorship

We derive a number of other variables from the RAE2008 data source. We capture the number of papers submitted, with a maximum of four per researcher. Researchers were able to be returned for the RAE with fewer than four if there were “individual staff circumstances” such as maternity leave, part-time working, or if the person was classified as an early career researcher. We also utilise information on co-authorship that allows us to determine whether an individual co-

²⁰ Journal articles accounted for over 90% of outputs submitted to the Business and Management Sub-Panel in RAE 2008.

authored within or between institutions, and include the number of co-authors for a piece of work, with a maximum of three.

3.3. Gender and rank

We conducted a web search to establish the gender and employment rank (Professor, Reader, Senior Lecturer, Lecturer, and a composite of other ranks) of all individuals submitted. We were able to capture all but 47 individuals' genders and 299 individuals' ranks from a total sample of 4,280 individuals (1.1% and 6.9% of the sample respectively). Gender was coded as a fixed binary variable. The data collection was undertaken between December 2011 and January 2012, so the variable probably represents an upwardly biased picture of what the employment ranks would have been in 2008, when all the research quality indicators were compiled, since in a competitive area such as business and management, it seems likely that many academics would have been promoted.

3.4 Control variables

As well as data on the quality sub-profiles for each institution, we extracted information on the number of submitted faculty (Category A), and the number of listed faculty in each of categories B to D,²¹ the headcount of research assistants and research officers (which we summed for our analysis), the number of PhD studentships, and the total value of grant income received by the submitting unit over the assessment period. We also determined whether the university was a member of the Russell Group.²² The numbers of research assistants, PhD studentships and grant income are all normalised by the Category A faculty total.

4. Results

²¹ Category A staff are defined as academic staff in post and on the payroll of the submitting institution on the census date. Category B staff are those who held a contract during the pre-RAE period 2001–07 but who left. Category C staff are independent investigators whose research is focused within the submitting department, while Category D staff are those who were Category C staff at some point during 2001–20–07, but who left before the “census date” (which was 31 October 2007).

²² The Russell Group is an association, formed in 1994, of 20 of the UK's largest (and in general, most prestigious) research-intensive universities that had medical schools. The size of this grouping has altered over time. We include institutions that were Russell Group members on the RAE census date in 2008.

Table 1 summarises the mean differences in output quality measures based on journal ratings between men and women for each sub-discipline (raw differences) and in the participation of women in terms of numbers entered into the RAE. The classification breakdown comes from the ABS list and is generally uncontentious. Several characteristics emerge from this descriptive data. Women are broadly represented across the field of business and management although there are marked differences in the choice of the subject area in which they work.

<INSERT TABLE 1 ABOUT HERE>

Table 1 also suggests that there are differences in journal ratings between the papers submitted to the RAE by women and men. The average differences between women and men provided for each indicator show that women, on average across all fields, have lower scores. However, there are a number of instances of sub-disciplines where women have higher scores with seven ABS sub-disciplines and eight ISI categories respectively but only one of the fifteen areas was included in the FT lists. There are also suggestions that not all indicators are similar. There is more variation in the impact factor measure than the ABS scores provides. The degree of difference is greater in the context of the “elite” (by virtue that it includes so few journals) FT listings. There are also instances where women have higher mean scores in one area than another. A key example, given that it is where female participation rates are highest, is human resource management (HRM). For this sub-field, the average ABS score is lower while the impact factor score is slightly higher for women. However, there is a marked increase in the difference between the average number of papers that were submitted by men and women in the HRM area when comparing those included in the FT 40 versus 45 lists. This change reflected the removal of *International Journal of Human Resource Management* from the FT40 listing since that journal was the most important outlet for women, accounting for around 20% of publications submitted to the RAE 2008 from the FT list.

4.1 RAE panel assessments

To assess whether the gender compositions of submissions directly influenced the results of the RAE 2008,²³ we derive a set of quality indicators and other factors that may influence the RAE outcomes. We estimate a rich specification using OLS that includes the proportion of women in an institution, gender, and a limited set of controls in the form of academic ranks: Professor, Reader, Senior Lecturer/Associate Professor, Lecturer/Assistant Professor, and another which incorporates a heterogeneous group of ranks including Senior Researchers, Deans, etc. (this group is used as the reference category in subsequent regressions) and a measure of output quality. *A priori*, we expect that professors will have a premium, as they are more experienced and established faculty members but it is less clear, given the heterogeneous nature of the reference group, how other ranks will compare. As our focus is on gender, this is not a critical issue. Specifically, we incorporate measures of the distribution of ABS and impact measures, number of papers submitted, number of co-authors, whether there was an external co-author, the numbers of Category A, B, C and D staff, whether or not the university was a member of the Russell Group, the amount of grant income secured, the number of research assistants employed, and the total number of PhD studentships awarded, as well as a full set of subject sub-discipline dummy variables.

In Table 2 we present the findings. The four columns containing data differ in the metric used to explain the RAE scores: baseline without metrics (column 2), the ABS rating (column 3), the impact factor (column 4) and the number of papers in FT journals normalised by the number of Category A faculty submitted (the FT40 list in column 5 and the FT45 in column 6). Given that the dependent variable is the same within each panel of Table 2, we can compare the R^2 values as indicators of the relative fit of the four quality measures used. On that basis, the ABS list stands out as having the greatest explanatory power, with the FT measures having by far the least.

²³ The RAE score combines three separate dimensions (outputs, research environment and esteem) through which the research exercise was conducted. Our interest and focus in this paper is on research outputs.

Turning to the key variables of interest, the proportion of men is not a significant factor regardless of the indicator of research productivity used, thereby refuting hypothesis 1. Hypothesis 2 is also not supported as there is no evidence that higher amounts of co-authorship external or internal to the institution influenced the results of the exercise. Finally, institutions with higher proportions of women who had “individual staff circumstances” did not fare any worse than those who did not, refuting hypothesis 3. We find in the absence of any quality measure (i.e. regression 1 in Table 2) that concentration of research in some sub-fields had significantly more positive effects on the RAE output score than others, namely, business history, psychology and social science, while others had significantly lower effects, namely, international business and area studies, and sector studies. Where the ABS list is used to capture quality there is no indication that particular sub-disciplines are favoured by the panel. However, business history and general social science are less well captured by the ISI and FT measures of research productivity, while international business and area and sector studies are over represented.²⁴ Therefore we find modest support for hypothesis 4.

<INSERT TABLE 2 ABOUT HERE>

Turning to the control variables, we find that a greater dispersion in the citation impact factor had a positive impact on outcomes. We did not find, unlike previous studies, that there is a Russell Group “halo” effect at work.

In summary, having controlled for a multitude of factors, including sub-disciplines, we find that gender had no direct influence on the outcomes of the RAE but that controlling for output quality is a critical factor. However, gender influences may occur through the journal quality measures themselves, and it is to the individual level of analysis that we now turn.

4.2 Determinants of an individual output’s rating

²⁴ Given the sample size we cannot provide interactions since this would mean that we lose a further 22 degrees of freedom. We did however examine interactions where there were significant sub-discipline results but did not find evidence of within-subject effects.

The literature does not provide a consistent view that women are affected differently by the use of journal ratings. Whether this is the case in business and management is, as far as we are aware, untested. We examine whether gender influences the four quality measures of interest, which now constitute the dependent variables utilising a specification similar to that used above when we examined the RAE scores directly.

There are important differences between the preceding regressions and those presented below. First, the entities employed in the regression are now the individual outputs submitted rather than aggregates for each institution. Ideally, this part of the study would also use ratings from the RAE2008, since although the journal quality measures (ABS, impact factor, FT) are able to reasonably model the outcomes of the RAE on aggregate where errors cancel out to some extent, the same may not be true for the evaluations of individual outputs. However, it is clearly not possible to use scores derived from the RAE results since the exercise was intended to evaluate the quality of the research of a submitting unit at the aggregate level, and consequently, the ratings awarded for individual pieces of work were not made public. Second, in terms of explanatory variables, we include a full set of institution dummies but we exclude the range variables as it would be tautological to include these. Given the differing natures of each measure, we adopt estimation strategies appropriate to each data form: a probit regression for the ordinal ABS measure, a generalised least squares (GLS) estimator for the continuous impact factor measure, and a probit estimator for the two *Financial Times* ratings. The gender variable now takes the value 1 if the listed author of the paper is a man and 0 for a woman.

The results are presented in Table 3. To allow for correlations in the measures within an institution, we employ clustering by institution. We find no gender significant factor regardless of the indicator of research productivity used, thereby refuting hypothesis 1.1. We do find, however, robust evidence that women co-authoring externally have a positive influence on all four quality measures, supporting hypothesis 2.1. Note that this indicates that women who co-author do better

than those who do not. However, an interesting aspect of this finding is that the positive co-authorship for women does not translate positively for men.

<INSERT TABLE 3 ABOUT HERE>

We also find that for the two measures more widely used outside the UK – the FT list and ISI impact factors – women who have “individual circumstances” are disadvantaged. This supports hypothesis 3.1 and is consistent with some of the literature arguing that gender has an effect on some bibliometrics but not others (HEFCE, 2011). Since “individual circumstances” include early career researchers, we cannot be certain that it is not young women, rather than women with other circumstances such as part-time work or maternity leave. However, we do measure rank directly, with lower ranks generally capturing early career academics. We also find that there is no penalty from having fewer than four outputs for men, who in the main are not as significantly engaged in child rearing or part-time work. Given the far greater tendency for men to work full time, and the disproportionate effects of early child rearing on women, we can be confident that these factors are at work. The findings in Table 3 also imply very different average measures of research quality across sub-disciplines. For example, when we compare the ABS and impact factor ratings, there are three sub-disciplines – HRM and employment, sector studies and strategy –where there are significant but opposing signs on the coefficients. In other words, the two measures are providing opposing indications of whether the average work in that discipline is rated significantly differently from that in the reference category, accounting. In other instances, such as ethics and governance, the parameter estimates for the ABS score regressions were insignificantly different from the accounting reference group, while those for the impact factor measure were significantly determined at the 5% level or better. Given the uneven participation rates across genders, these findings potentially have important implications for women that will be discussed in the next section.

4.3 Gender differences within sub-disciplines

A further mechanism by which gender may influence the measured quality of the journals in which academics publish is through subject specialisms. Women are better represented in some subject areas than others, as was illustrated in Table 1. We now run a total of 76 individual estimations (24 for each of the ABS and impact factors and 14 for each of the FT40 and 45 lists)²⁵ that follow identical specifications to those used on the full datasets to ascertain whether or not there is evidence of gender effects within sub-disciplines. We examine each sub-discipline separately, removing all of the outputs from the other sub-disciplines from that regression. The results are presented in Table 4.²⁶ We find limited evidence of gender impacts within subject areas. There is only one positive (indicating that men on average are scoring higher) and significant coefficient for gender on any of the quality measures relating to management development and education. There is an indication that women earn on average *higher* ratings in marketing for the ABS list measure. Consistent with earlier findings, the FT measures differ. The FT40 has three positive and well determined sub-discipline coefficients of the 14 areas: accountancy, HRM and strategy, while there is a negative coefficient in relation to information management. The alteration of the FT list has increased the proportion of women's research published in these journals; however, there are only two sub-disciplines where men have significantly higher percentages of their papers in FT45 listed journals whereas there were four such sub-disciplines for the previous incarnation. We find where there are gender impacts in the majority of cases these benefit men. Our results suggest tentative support for hypothesis 4.1.

<INSERT TABLE 4 ABOUT HERE>

5. Discussion

While we find little evidence of gender effects in RAE panel evaluations, the findings in relation to journal ratings lists have important implications for women in terms of their choice of

²⁵Only 14 are included for the FT measures because 10 of the sub-disciplines contain no journals in the FT lists and therefore there are no data for those subject areas.

²⁶As a referee has cautioned us, given the large number of parameter estimates, it is possible that some estimates may be significant by chance alone. However, this is partially mitigated in that findings on key variables are robustly determined, in many cases being significant at the 1% level or better. We have also experimented with a variety of simplified specifications with the results being similar to those provided in Table 4.

sub-discipline. Changes in the composition of the FT list have had significant effects on particular sub-disciplines, with organisation studies shifting from having a negative coefficient (women doing better) in the FT40 regression to having a positive coefficient (men doing better) in the FT45. General management shifts from having an insignificant effect to having a positive coefficient while HRM and employment has a negative shift. Given the small number of journals involved, it is straightforward to appreciate that it is the inclusion in journal ratings lists for specific journals, and the ability/willingness of these journals to accept work by UK scholars that drive these changes between the two vintages of FT listing. For example, both the *Journal of Management Studies* and *Organisation Science*, two of the most important outlets for UK authors in terms of publications in RAE 2008, were to appear in the FT45, while the *International Journal of Human Resource Management*, the journal with the highest number of articles submitted to the RAE 2008 from the FT40 list, was subsequently deleted. Given the higher participation rates by women in HRM, this had a damaging effect upon their individual ratings.

In addition, our findings show that gender differences in scientific performance can partly be explained by women's ability to network and/or their individual circumstances. The significance of research networks (Fletcher, 2007; Haynes and Fearfull, 2008; McDowell et al., 2005; Pezzoni et al., 2012) for women is supported by our findings that, although the number of co-authors *per se* is not important, there is a positive and robust relationship between research quality and co-authoring with researchers outside the institution. This may explain why women who have individual staff circumstances publish in journals receiving lower impact factors. The ability to develop research networks, particularly outside the institution, is dependent upon the mobility of staff, such as being able to travel to conferences and stay away from home. Women with responsibilities for children will find it more difficult to arrange childcare and even during the working day may have their time curtailed by nursery and school runs. However, Jordan et al. (2008) provide evidence that women may also suffer from gender selection bias, as their study of accounting faculty show men tend towards male co-authors while women tend toward female

collaborators. As many academic institutions are under-represented by women at higher academic levels, this gender selection bias may put women at a disadvantage for finding suitable research partners, thereby explaining their lower publication productivity.

Women classified as having individual circumstances perform less well than those without by the FT40 but not directly so for the other three journal ratings lists, which calls into question the robustness of such lists to assess individuals fairly. Aside from the direct effect of gender, we find a host of significant determinants of research quality at the individual output level. The number of outputs that a researcher returned for the RAE is a well determined factor in three of the four measures. This is of concern in relation to gender as 14.1% of women, compared with just 8.2% of men, submitted fewer than four outputs. Thus the effect of mitigating circumstances, such as maternity leave, might not only reduce the number of available pieces, but also their quality. Or it could be that such researchers simply had less to choose from.

6. Conclusion and implications

In this paper, we have provided a framework to evaluate possible gender effects in the evaluation of research when either assessment panels or journal ratings are utilised. While we did not find evidence of gender effects where the assessment of business and management was concerned, we did find evidence that women may be disadvantaged by their sub-discipline specialisations when their research is evaluated using journal ratings lists. We have demonstrated that the decision by HEFCE to limit their use in the REF was a sensible one. Furthermore, our findings demonstrate that a lack of networks and individual circumstances are important explanatory variables in differential quality outcomes for men and women in Business and Management.

So how might women help themselves? As career progress in academia is often explained by affiliation to important public research organisations and social ties with senior members of the discipline (Pezzoni et al, 2012; Finch, 2003), it behoves women to promote both their productivity and visibility (Hunter and Leahey, 2010). Individual career planning in these

respects should begin early, with women embedding themselves in scholarly communities (Leahey et al, 2008) and gaining the support of competent mentors and role models to build confidence (Fletcher 2007). Additionally, having a coherent collaborator choice strategy (either instrumental or experience based) has been shown to predict the number of collaborators for both men and women (Bozeman and Gaughan, 2011) and women should be advised to build networks with men rather than just women. As specialisation has been shown to positively effect the level of productivity for women (Leahey et al, 2008), women should plan their career with the extent of specialisation in mind, giving consideration to publication outlets.

It is worth considering some broader implications for research policy and areas where future research is needed. One potentially important dimension is a legal one for the higher education sector, given that legislation requires the public sector to take proactive steps to promote equality rather than simply prevent discrimination. The possibility of gender effects in the use of journal ratings lists implies that they should be used with caution by institutions, although such lists are often popular among business school Deans and Directors of Research as a way to rationalise decisions on tenure, hiring, promotion and pay. Indeed, given that the range of sub-fields in business and management is so wide, those responsible for making decisions at the institution-wide level, and even those responsible for evaluating the quality of research at the school or departmental level, will rarely have the skills or time to be able to evaluate quality accurately in all areas. The impact of the diversity of sub-fields that collectively comprise the business and management area on hiring and promotion decisions may be exacerbated by the typical UK academic interview panel. Unlike in other countries, such as the United States, UK interview panels comprise senior university staff members from outside the discipline and the Dean but may include few subject specialists. Our findings indicate that interview panels require a broader constitution of knowledgeable specialists in the candidates' sub-field(s) rather than the typical mix of leaders and lay people. Alternatively, panels could solicit the evaluations of such

specialists, who commit to actually reading the work, prior to interviews instead of relying on journal ratings lists.

A related issue is the impact on individuals via pay and promotion. While measures of quality have real impacts on these, work in this area is not well developed (Barbezat, 1987; Blackaby et al., 2005), and more recent research has highlighted a pay gap between men and women (Walker et al., 2010; Ward, 2001b). Given the pace at which metrics have been adopted by nation states and institutions, the issue is a pressing one. Our findings indicate that as with hiring policies, pay and promotions decisions should not rely upon journal ratings lists but employ evaluations by specialists within the specific sub-discipline. Future research could assess the relationship between promotional policies, including their use of research evaluation measures, and the pay or promotional outcomes between men and women in UK universities.

The indirect effects of journal ratings lists on women in terms of their networking abilities, with individual circumstances and certain sub-discipline specialisation, suggest a requirement for research quality assessments within institutions to dispense with the use of journal ratings lists (Nightingale and Scott, 2007) and promote the use of panel-based processes instead. It has been suggested that panel-based processes should be supplemented with sophisticated indicators measuring actual citation impacts at the paper level (Butler, 2007; Moed, 2007) including those that differentiate between sub-fields (Moed, 2005).

Our findings also have implications for other disciplines in the social sciences and humanities where academics tend to publish in books rather than journals. This is to eschew the temptation to develop “quality ratings lists” of publishers where merely publishing with one imprint rather than another would result in a higher rating being awarded, thereby encouraging gamesmanship strategies, but rather to conduct quality assessment exercises via panel processes. Peer evaluation processes are a key characteristic of the UK approach to research assessment and are seen as maintaining the legitimacy of the RAE (Broadbent, 2010).

While the present study focuses on gender, our findings can also be viewed through a wider lens. Recent work has shown journal ratings are biased against inter-disciplinary research (Rafols et al., 2012) and emerging disciplines and publications (Donovan, 2007). Taken together these findings suggest that a shift towards the “objective assessment” of research using journal lists or other crude quantitative measures may tend to reduce diversity and in some cases may blunt women’s career prospects while impeding intellectual discovery and the development of knowledge-based economies (Donovan, 2007; Northcott and Linacre, 2010; Piercy, 2000).

References

- Agrawal, V.K., Agrawal, V., Rungtusanatham, M., 2011. Theoretical and interpretation challenges to using the author affiliation index method to rank journals. *Production and Operations Management* 20(2), 280–300.
- Aksnes, D.W., Rorstad, K., Piro, F., Gunnar, S., 2011. Are female researchers less cited? A large-scale study of Norwegian scientists. *Journal of the American Society for Information Science and Technology* 62, 628–36.
- Barbezat, D.A., 1987. Salary differentials by sex in the academic labour market. *Journal of Human Resources* 22, 422–8.
- Beattie, V., Goodacre, A., 2012. Publication records of accounting and finance faculty promoted to professor: Evidence from the UK. *Accounting and Business Research* 42(2), 197–231.
- Blackaby, D., Booth, A., Frank, J., 2005. Outside offers and the gender pay gap: Empirical evidence from the UK academic labour market. *The Economic Journal* 115, F81–F107.
- Bozeman, B., Gaughan, M., 2011. How do men and women differ in research collaborations? Analysis of the collaborative motives and strategies of academic researchers. *Research Policy* 40(10), 1393–402.
- Broadbent, J., 2010. The UK Research Assessment Exercise: Performance measurement and resource allocation. *Australian Accounting Review* 20, 14–23.
- Browning, M., 1992. Children and household economic behavior. *Journal of Economic Literature* 30, 1434–75.
- Budig, M.J., England P., 2001. The wage penalty for motherhood. *American Sociological Review* 66, 204–25.
- Butler, L., 2007. Assessing university research: A plea for a balanced approach. *Science and Public Policy* 34(8), 565–74.
- Collini, S., 2008. *Absent Minds: Intellectuals in Britain*. Oxford University Press, Oxford.
- Colyvas, J., 2012. Performance metrics as formal structures and through the lens of social mechanisms: When do they work and how do they influence? *American Journal of Education* 118(2), 167–97.
- Coupé, T., Ginsburgh, V., Noury A., 2010. Are leading papers of better quality? Evidence from a natural experiment. *Oxford Economic Papers* 62, 1–11.
- Davenport, E., Snyder, H., 1995. Who cites women? Whom do women cite? An exploration of gender and scholarly citation in sociology. *Journal of Documentation* 51, 404–10.
- Department for Business Innovation and Skills, 2012, *Annual Innovation Report: Innovation, Growth, and Research*.
- Donovan, C., 2007. The qualitative future of research evaluation. *Science and Public Policy* 34(8), 585–97.

Dwyer, P.D., 1994. Gender differences in the scholarly activities of accounting academics: An empirical investigation. *Issues in Accounting Education* 9, 231–47.

Easton, G., Easton, D.M., 2003. Marketing journals and the Research Assessment Exercise. *Journal of Marketing Management* 19, 5–24.

European Commission, 2005. Realising the European higher education area: Achieving the goals. Conference of European Higher Education Ministers Contribution of the European Commission Bergen, 19/20 May 2005, Brussels.

Findlay, A., Sparks, L., 2010. The UK Research Assessment Exercise (RAE 2008) and retail research output. *International Review of Retail, Distribution & Consumer Research* 20(5), 583–95.

Fletcher, C., 2007. Passing the buck: Gender and management of research production in UK higher education: Management perspectives from a case study. *Equal Opportunities International* 26, 269–86.

Giles, M.W., Garand, J.C., 2007. Ranking political science journals: Reputational and citational approaches. *PS: Political Science & Politics* 40(4), 741–51.

Ginther, D.K., Kahn, S., 2004. Women in economics: Moving up or falling off the academic career ladder? *Journal of Economic Perspectives* 18, 193–214.

Haynes, K., Fearfull, A., (2008). Exploring ourselves: Exploiting and resisting gendered identities of women academics in accounting and management. *Pacific Accounting Review* 20, 185–204.

HEFCE, 2011. Analysis of data from the pilot exercise to develop bibliometric indicators for the REF: The effect of using normalised citation scores for particular staff characteristics. Issues Paper, February 2011/03.

Hoepner, A.G.F., Unerman, J., 2012. Explicit and implicit subject biases in the ABS Journal Quality Guide. *Accounting Education* 21, 3–15.

Hunter, L.A., Leahey, E., 2010. Parenting and research productivity: New evidence and methods. *Social Studies of Science* 40, 433–51.

Jacobs, J.A., Winslow, S.E., 2004. The academic life course, time pressures and gender inequality. *Community, Work and Family* 7(2), 143–61.

Jordan, C.E., Clark, S.J., Vann, C.E., 2008. Do gender differences exist in the publication productivity of accounting faculty? *Journal of Applied Business Research* 24, 77–85.

Kelly, A., Morris, H., Harvey, C., 2009. Modelling the outcome of the UK Business and Management Studies RAE 2008 with reference to the ABS Journal Quality Guide (downloaded on 29 April 2011 from http://www.the-abs.org.uk/files/RAE2008_ABS2009_final.pdf).

- Knights, D., Richards, W., 2003. Sex discrimination in UK academia. *Gender, Work and Organisation* 10, 213–38.
- Leahey, E., Crockett, J. L., Hunter, L.A., 2008. Gendered academic careers: Specializing for success? *Social Forces* 86(3), 1273–309.
- London, B., Rattan, D.G., Romero-Canyas, R., Tyson, D., 2012. Gender-based rejection sensitivity and academic self-silencing in women. *Journal of Personality & Social Psychology* 5, 961–79.
- Long, S.J., Allison, P.D., McGinnis, R., 1993. Rank advancement in academic careers: Sex differences and the effects of productivity. *American Sociological Review* 58(5), 703–22.
- Macdonald, S., Kam, J., 2007a. A ring a ring o’ roses: Quality journals and gamesmanship in management studies. *Journal of Management Studies* 44, 640–55.
- Macdonald, S., Kam, J., 2007b. Quality journals and gamesmanship in management studies. *Journal of Information Science* 33, 702–17.
- Maranto, C.L., Griffin, A.E.C., 2011. The antecedents of a ‘chilly climate’ for women faculty in higher education. *Human Relations* 64, 139–59.
- Mayer, A.L., Tikka, P.M., 2008. Family-friendly policies and gender bias in academia. *Journal of Higher Education Policy & Management* 30(4), 363–74.
- McDowell, J.M., Singell, L.D., Stater, M., 2005. Two to tango? Gender differences in the decisions to publish and co-author. *Economic Inquiry* 44, 153–68.
- McElhinny, B., Hols, M., Holtzkenner, J., Unger, S., Hicks, C., 2003. Gender, publication and citation in sociolinguistics and linguistic anthropology: The construction of a scholarly canon. *Language in Society* 32, 299–328.
- Miller, C.C., Glick, W. H., Cardinal, L.B., 2005. The allocation of prestigious positions in organizational science: Accumulative advantage, sponsored mobility and contest mobility. *Journal of Organizational Behaviour* 26(5), 489–516.
- Millimet, D.L., 2000. The impact of children on wages, job tenure, and the division of household labour. *Economic Journal* 110, 139–57.
- Moed, H.F., 2005. *Citation Analysis in Research Evaluations*. Springer, Dordrecht, Netherlands.
- Moed, H.F., 2007. The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy* 34(8), 575–83.

- Morris, H., Harvey, C., Kelly, A., 2009. Journal rankings and the ABS Journal Quality Guide. *Management Decision* 47, 1441–51.
- Morris, H., Harvey, C., Kelly, A., Rowlinson, M., 2011. Food for thought? A rejoinder on peer-review and RAE2008 evidence. *Accounting Education* 20, 561–73.
- Nightingale, P., Scott, A., 2007. Peer review and the relevance gap: Ten suggestions for policy-makers. *Science and Public Policy* 34, 543–53.
- Northcott, D., Linacre, S., 2010. Producing spaces for academic discourse: The impact of Research Assessment Exercises and journal quality rankings. *Australian Accounting Review* 20(1), 38–54.
- Oswald, A.J., 2007. An examination of the reliability of prestigious scholarly journals: Evidence and implications for decision-makers. *Economica* 74, 21–31.
- Pezzoni, M., Sterzi, V., Lissoni, F., 2012. Career progress in centralized academic systems: Social capital and institutions in France and Italy. *Research Policy* 41, 704–19.
- Piercy, N., 2000. Why it is fundamentally stupid for a business school to try to improve its Research Assessment Score. *European Journal of Marketing* 34, 27–36.
- Probert, B., 2005. 'I just couldn't fit it in': Gender and unequal outcomes in academic careers. *Gender, Work & Organization* 12, 50–72.
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A., 2012. How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy* 41, 1262–82.
- Rama, D., Raghunandan, K., Logan, L.B., Barkman, B.V., 1997. Gender differences in publications by promoted faculty. *Issues in Accounting Education* 12(2), 353–65.
- Reinstein, A., Calderon, T.G., 2006. Examining accounting department's rankings of the quality of accounting journals. *Critical Perspectives on Accounting*, 17, 457–90.
- Sealy, R., Singh V., 2010. The importance of role models and demographic context for senior women's work identity development. *International Journal of Management Reviews* 12, 284–300.
- Simonsen, M., Skipper, L., 2008. An empirical assessment of effects of parenthood on wages. *Advances in Econometrics*, 21, 359–80.
- Stewart, J., 2005. The UK Research Assessment Exercise. *Human Resource Development International* 8, 489–94.

- Taylor, J., 2011a. The assessment of research quality in UK universities: Peer review or metrics? *British Journal of Management*, 22, 202–17.
- Taylor, J., 2011b. How biased were the peer review panels in the UK's 2008 Research Assessment Exercise? Mimeo, Department of Economics, Lancaster University.
- Times Higher Education, 2009. Unequal opportunities in final RAE. 17 September 2009 (downloaded on 26 July 2013 from <http://www.timeshighereducation.co.uk/story.asp?storycode=408202>).
- Van Den Brink, M., Brouns, M., Waslander, S., 2006. Does excellence have a gender? *Employee Relations* 28(6), 523–39.
- Voss, C., 2010. Reflections on the state of OM. *POMS Chronicle* 17(1), 11–12.
- Waldfogel J., 1998a. Understanding the “family gap” in pay for women with children. *Journal of Economic Perspectives* 12, 137–56.
- Waldfogel J., 1998b. The family gap for young women in the United States and Britain: Can maternity leave make a difference? *Journal of Labor Economics* 16, 505–45.
- Walker, J.T., Vignoles, A., Collins M., 2010. Higher education academic salaries in the UK. *Oxford Economic Papers* 62, 12–35.
- Ward, K.B., Gast, J., Grant, L., 1992. Visibility and dissemination of women's and men's sociological scholarship. *Social Problems* 39, 291–98.
- Ward, M., 2001a. Gender and promotion in the academic profession. *Scottish Journal of Political Economy* 48, 283–302.
- Ward, M., 2001b. The gender salary gap in British academia. *Applied Economics* 33, 1669–81.
- Wolfinger, N.H., Mason, M.A., Goulden, M., 2009. Stay in the game: Gender, family formation and alternative trajectories in the academic life course. *Social Forces* 87(3), 1591–621.

Table 1: Mean differences in journal ratings for sub-disciplines (raw differences) and in participation of women in those sub-disciplines

	ABS			Impact Factor			FT40			FT45			Women	Men	Proportion
	Women	Men	% diff	Women	Men	% diff	Women	Men	% diff	Women	Men	% diff	No.	No.	Women
ACCOUNT	3.0	3.0	-0.3	1.37	1.31	4.28	0.1	0.2	-22.0	0.15	0.21	-27.5	115	433	20.99
BUS HIST	3.0	2.9	4.5	0.54	0.60	-9.31							24	62	27.91
ECON	2.6	2.7	-3.3	0.86	0.90	-5.24	0.0	0.0	-20.7	0.03	0.03	-15.4	175	899	16.29
ENT-SMBUS	2.8	2.8	-2.5	1.29	1.37	-5.61	0.2	0.2	-9.9	0.21	0.23	-9.9	79	206	27.72
ETH-GOV	2.7	2.6	7.0	0.95	0.96	-1.07	0.5	0.4	34.2	0.50	0.37	34.2	59	113	34.30
FINANCE	2.8	2.8	-0.9	1.08	1.21	-10.83	0.1	0.1	-20.6	0.11	0.13	-20.6	112	668	14.36
GEN MAN	3.1	3.2	-3.1	1.66	1.88	-11.64	0.1	0.2	-58.8	0.35	0.44	-21.0	240	623	27.81
HRM&EMP	2.9	3.1	-5.9	0.88	0.86	2.35	0.1	0.2	-26.4	0.01	0.01	-32.2	296	505	36.95
IB&AREA	2.5	2.9	-14.3	1.37	1.55	-11.77	0.2	0.4	-42.3	0.16	0.31	-47.4	40	202	16.53
INFO MAN	2.7	2.7	1.3	1.35	1.38	-1.93	0.0	0.0	25.8	0.04	0.04	25.8	103	253	28.93
INNOV	2.4	2.3	6.3	1.03	0.98	5.38							31	89	25.83
LAW	2.3	2.3	2.4	1.62	0.68	138.94							10	22	31.25
MGDEV&ED	1.9	1.8	3.7	1.16	1.79	-35.28							54	111	32.73
MKT	2.5	2.5	0.2	0.87	1.04	-16.13	0.0	0.1	-40.9	0.04	0.07	-40.9	287	686	29.50
OPS&TECH	2.5	2.6	-4.8	0.83	0.93	-10.67	0.0	0.1	-60.4	0.02	0.07	-68.3	106	504	17.38
OR&MANSCI	3.0	3.0	-0.4	0.99	1.07	-7.41	0.1	0.1	-38.7	0.04	0.06	-29.4	108	416	20.61
ORG STUD	3.4	3.6	-6.3	1.91	1.87	2.17	0.0	0.1	-56.5	0.21	0.34	-37.9	172	402	29.97
PSYCH	2.9	3.3	-10.9	1.99	2.02	-1.37	0.1	0.3	-59.2	0.11	0.26	-59.2	59	130	31.22
PUB SEC	2.4	2.5	-2.3	0.82	0.79	4.03							83	254	24.63
SECTOR	1.8	2.0	-10.0	1.19	1.43	-16.92							63	149	29.72
SOC SCI	3.1	3.2	-4.1	1.52	1.51	0.61							171	526	24.53
STRAT	2.4	2.8	-14.4	1.57	1.81	-13.16	0.1	0.3	-79.9	0.05	0.27	-79.9	42	152	21.65
TOUR-HOSP	2.2	2.7	-18.1	0.88	0.88	0.52							67	168	28.51
AVERAGE DIFF			-3.31			-12.64			-31.76			-28.64			26.06

Notes: Categories follow ABS breakdown. Law is excluded from the average difference in the FT measures due to the small number of participants.

Table 2: Determinants of RAE panel results

Research Outputs		1		2		3		4		5	
		coeff	t	coeff	t	coeff	t	coeff	t	coeff	t
ABS score				0.817	(7.91)						
Impact factor						0.422	(3.37)				
FT40								30.052	(2.17)		
FT45										5.848	(0.38)
ABS range				-0.063	(1.41)						
Impact factor range						0.228	(3.22)				
Gender		-1.570	(0.67)	-0.335	(0.16)	0.237	(0.11)	-0.726	(0.27)	-1.353	(0.47)
Rank	Professor	1.229	(3.83)	0.346	(1.81)	0.779	(3.02)	1.083	(3.47)	1.214	(3.78)
	Reader	1.597	(3.51)	0.548	(1.51)	1.357	(3.16)	1.727	(3.74)	1.648	(3.22)
	Senior Lecturer	0.813	(2.05)	0.428	(1.78)	0.808	(2.58)	1.061	(2.76)	0.840	(2.18)
	Lecturer	0.744	(2.26)	0.150	(0.68)	0.555	(2.11)	0.854	(2.70)	0.778	(2.34)
No. coauthor(s) (woman)		0.304	(0.86)	0.136	(0.52)	0.466	(1.29)	0.230	(0.57)	0.333	(0.86)
No. coauthor(s) (man)		-0.033	(0.07)	-0.327	(0.96)	-0.377	(0.92)	0.022	(0.05)	-0.047	(0.10)
External coauthor(s) (woman)		0.557	(0.80)	-0.172	(0.43)	-0.028	(0.04)	0.327	(0.49)	0.584	(0.81)
External coauthor(s) (man)		-0.651	(0.73)	-0.040	(0.07)	0.096	(0.12)	-0.540	(0.65)	-0.705	(0.78)
No. of papers (woman)		0.005	(0.01)	-0.103	(0.33)	0.054	(0.14)	0.151	(0.37)	0.060	(0.13)
No. of papers (man)		0.704	(1.24)	0.214	(0.44)	0.177	(0.32)	0.416	(0.60)	0.659	(0.93)
No. of category A staff		0.003	(1.85)	0.003	(2.80)	0.003	(2.11)	0.004	(2.53)	0.003	(1.79)
No. of category B staff		-0.001	(0.20)	-0.001	(0.46)	-0.002	(0.59)	0.000	(0.01)	0.000	(0.09)
No. of category C staff		-0.019	(0.59)	-0.006	(0.28)	-0.026	(1.07)	-0.026	(0.80)	-0.018	(0.54)
No. of category D staff		0.073	(2.26)	0.035	(1.58)	0.038	(1.44)	0.063	(2.05)	0.073	(2.29)
Russell Group		0.045	(0.67)	0.021	(0.48)	-0.008	(0.15)	0.044	(0.72)	0.045	(0.66)
Grant income		0.000	(1.02)	0.000	(0.53)	0.000	(0.13)	0.000	(0.39)	0.000	(0.75)
Studentships		0.000	(0.74)	0.000	(1.67)	0.000	(0.84)	-0.001	(1.14)	0.000	(0.80)
Research assistants		-0.005	(0.90)	-0.002	(0.61)	-0.001	(0.24)	-0.004	(0.70)	-0.004	(0.74)
BUS HIST		4.071	(2.90)	0.144	(0.12)	2.951	(2.33)	2.941	(2.06)	3.972	(2.71)
ECON		0.699	(1.01)	0.638	(1.10)	0.640	(1.27)	0.710	(1.12)	0.747	(1.10)
ENT-SMBUS		-0.212	(0.28)	-0.278	(0.51)	-0.721	(1.62)	-0.509	(0.72)	-0.223	(0.30)
ETH-GOV		0.091	(0.05)	1.228	(1.21)	0.686	(0.46)	-0.578	(0.32)	0.020	(0.01)
FINANCE		0.219	(0.29)	-0.209	(0.33)	0.140	(0.20)	-0.117	(0.16)	0.239	(0.30)
GEN MAN		1.523	(1.79)	-0.051	(0.08)	0.272	(0.46)	1.257	(1.51)	1.466	(1.61)
HRM&EMP		0.373	(0.54)	-0.448	(0.72)	0.433	(0.78)	-0.017	(0.02)	0.386	(0.56)
IB&AREA		-2.335	(2.27)	-0.369	(0.50)	-2.181	(2.88)	-2.593	(2.70)	-2.323	(2.25)
INFO MAN		0.643	(0.66)	0.010	(0.02)	0.532	(0.73)	0.498	(0.55)	0.687	(0.74)
INNOV		-2.914	(1.82)	-0.481	(0.28)	-3.177	(2.02)	-2.622	(1.71)	-2.765	(1.59)
LAW		-2.252	(0.91)	-2.255	(1.39)	-4.184	(1.72)	-2.663	(1.04)	-1.941	(0.67)
MGDEV&ED		-1.393	(1.62)	0.558	(0.75)	0.234	(0.35)	-1.351	(1.74)	-1.278	(1.47)
MKT		0.160	(0.26)	0.522	(1.03)	0.428	(0.78)	0.035	(0.06)	0.200	(0.31)
OPS&TECH		0.388	(0.47)	0.502	(0.75)	0.716	(1.05)	0.343	(0.42)	0.461	(0.53)
OR&MANSCI		0.672	(0.76)	0.018	(0.03)	0.138	(0.17)	0.333	(0.40)	0.626	(0.68)
ORG STUD		0.662	(0.65)	-0.065	(0.08)	0.483	(0.53)	0.428	(0.44)	0.589	(0.55)
PSYCH		2.108	(2.00)	0.326	(0.53)	0.099	(0.10)	1.316	(1.17)	2.054	(1.91)
PUB SEC		-1.221	(1.58)	-0.179	(0.31)	-0.854	(1.44)	-1.545	(2.35)	-1.208	(1.64)
SECTOR		-2.651	(2.31)	-0.729	(1.14)	-2.338	(2.90)	-2.970	(2.62)	-2.661	(2.29)
SOC SCI		2.510	(3.52)	0.675	(1.16)	1.683	(2.74)	2.248	(3.40)	2.513	(3.56)
STRAT		0.053	(0.04)	-0.241	(0.25)	0.562	(0.57)	-1.066	(0.88)	-0.006	(0.01)
TOUR-HOSP		-0.215	(0.30)	-0.068	(0.11)	0.025	(0.04)	-0.309	(0.42)	-0.157	(0.21)
Constant		-0.289	(0.15)	-0.297	(0.20)	-0.689	(0.41)	-0.440	(0.24)	-0.542	(0.27)
N		89		89		89		89		89	
R ²		0.88		0.95		0.92		0.89		0.87	

Notes: the dependent variables are the RAE2008 grade point averages for outputs, the research environment and esteem in the first, second and third panels respectively. Each column in the table includes a different measure of journal quality. Gender measures the proportion of outputs in the submission produced by men; the rank variables denote the proportion of category A faculty that are within each category. N=89 observations, one for each of the institutions that made a return to the Business and Management Sub-Panel in RAE2008. t-statistics are reported and are derived using robust standard errors.

Table 3: Determinants of journal rating scores

		Ordered probit		GLM		Probit		Probit	
		ABS		Impact Factor		FT40		FT45	
		coeff	z	coeff	z	coeff	z	coeff	z
Gender		0.056	(0.35)	-0.099	(0.75)	-0.003	(0.07)	-0.012	(0.26)
No. of coauthor(s) (woman)		0.012	(0.42)	-0.002	(0.09)	-0.005	(0.56)	-0.010	(1.08)
No. of coauthor(s) (man)		-0.006	(0.17)	0.018	(0.67)	0.004	(0.39)	0.010	(1.03)
External coauthor(s) (woman)		0.226	(4.20)	0.105	(2.43)	0.057	(4.08)	0.042	(2.83)
External coauthor(s) (man)		-0.002	(0.04)	-0.029	(0.59)	-0.019	(1.15)	0.003	(0.16)
No. of papers (woman)		-0.051	(1.54)	-0.090	(3.13)	-0.018	(2.31)	-0.002	(0.23)
No. of papers (man)		-0.004	(0.10)	0.028	(0.83)	0.008	(0.83)	0.001	(0.11)
Rank	Professor	0.314	(5.74)	0.136	(3.66)	0.026	(1.88)	0.032	(1.98)
Ref. Other	Reader	0.094	(1.40)	-0.002	(0.04)	0.009	(0.49)	0.020	(0.95)
	Senior Lecturer	-0.136	(2.34)	-0.004	(0.11)	-0.007	(0.50)	-0.025	(1.56)
	Lecturer	-0.205	(3.21)	-0.084	(1.99)	-0.031	(2.21)	-0.041	(2.53)
No. of category A staff		0.039	(0.54)	-0.006	(3.48)	-0.002	(3.36)	-0.001	(1.81)
No. of category B staff		-0.004	(0.89)	0.018	(6.30)	0.005	(4.19)	0.005	(2.97)
No. of category C staff		0.035	(3.60)	-0.061	(1.85)	-0.007	(0.63)	-0.007	(0.41)
No. of category D staff		0.136	(1.26)	-0.004	(0.18)	0.002	(0.32)	0.006	(0.65)
Russell Group		0.164	(3.14)	0.094	(1.12)	0.067	(2.05)	0.084	(1.81)
Grant income		0.693	(2.90)	0.000	(5.34)	0.000	(3.97)	0.000	(3.20)
Studentships		0.000	(3.45)	0.002	(3.90)	0.000	(2.29)	0.000	(1.50)
Research assistants		-0.001	(0.56)	-0.033	(7.42)	-0.007	(4.85)	-0.008	(4.10)
Sub-discipline	BUS HIST	-0.010	(0.79)	-0.577	(0.35)				
Ref Accounting	ECON	0.144	(0.82)	-0.394	(8.56)	-0.089	(11.47)	-0.100	(12.07)
	ENT-SMBUS	-0.453	(8.04)	0.137	(2.71)	0.059	(2.95)	0.063	(2.87)
	ETH-GOV	0.076	(0.94)	-0.141	(2.13)	0.344	(9.83)	0.374	(9.85)
	FINANCE	-0.060	(0.74)	-0.171	(3.10)	-0.051	(5.92)	-0.060	(6.23)
	GEN MAN	-0.284	(4.94)	0.560	(9.61)	-0.019	(1.80)	0.153	(8.52)
	HRM&EMP	0.419	(6.18)	-0.284	(6.36)	0.038	(2.73)	-0.087	(8.06)
	IB&AREA	0.396	(6.60)	0.341	(4.70)	0.200	(7.86)	0.098	(4.32)
	INFO MAN	-0.071	(0.74)	0.116	(1.53)	-0.063	(6.40)	-0.076	(6.95)
	INNOV	-0.301	(4.76)	-0.213	(3.73)				
	LAW	-0.645	(6.13)	-0.403	(1.83)				
	MGDEV&ED	-0.565	(3.46)	0.265	(1.39)				
	MKT	-0.882	(0.08)	-0.226	(3.76)	-0.049	(5.42)	-0.060	(5.98)
	OPS&TECH	-0.342	(5.96)	-0.245	(5.39)	-0.052	(5.38)	-0.054	(4.80)
	OR&MANSCI	-0.360	(6.02)	-0.149	(3.30)	-0.040	(4.00)	-0.067	(6.41)
	ORG STUD	0.099	(1.75)	0.649	(1.59)	-0.060	(6.86)	0.076	(4.47)
	PSYCH	1.081	(3.73)	0.661	(8.06)	-0.010	(0.61)	-0.010	(0.54)
	PUB SEC	0.297	(2.43)	-0.295	(5.93)				
	SECTOR	-0.284	(4.35)	0.181	(2.44)				
	SOC SCI	-1.002	(9.65)	0.322	(6.73)				
	TOUR-HOSP	-0.227	(2.67)	-0.169	(3.27)				
Institutions fixed effects		YES		YES		YES		YES	
N		9,838		6,187		7,430		7,360	
Pseudo R ²		0.14		na		0.28		0.31	
Log likelihood		- 10,617		- 6,200		-2018.96		-2197.58	

Notes: the dependent variables are the individual quality scores for each journal output according to the ABS, impact factor or FT list measures, where for the latter the variable takes the value 1 if the output is in an FT-listed journal and zero otherwise. z-statistics are reported and are derived using robust standard errors.

Table 4: Gender coefficients from within sub-discipline regressions

	ABS coeff	z	Impact Factor coeff	z	FT 40 coeff	z	FT 45 coeff	z
ACCOUNT	0.178	(1.13)	0.042	(0.34)	0.151	(2.30)	0.174	(2.44)
BUS HIST	-0.580	(1.22)	0.026	(0.66)				
ECON	-0.005	(0.07)	-0.061	(0.93)	-0.029	(0.55)	-0.077	(1.17)
ENT-SMBUS	0.065	(0.26)	0.038	(0.49)	0.029	(0.22)	0.029	(0.22)
ETH-GOV	-0.232	(1.52)	0.063	(0.56)	-0.203	(1.08)	-0.203	(1.08)
FINANCE	-0.035	(0.26)	0.035	(0.32)	0.055	(0.98)	0.055	(0.98)
GEN MAN	-0.071	(0.85)	0.028	(0.29)	0.067	(1.72)	0.003	(0.07)
HRM&EMP	0.181	(1.85)	-0.064	(1.86)	0.091	(2.64)	-0.006	(0.08)
IB&AREA	0.067	(0.21)	-0.019	(0.10)	0.025	(0.17)	0.028	(0.23)
INFO MAN	-0.344	(1.61)	-0.190	(0.97)	-0.780	(3.63)	-0.780	(3.63)
INNOV	-0.590	(1.12)	-0.081	(0.48)				
MGDEV&ED	0.161	(1.00)	2.132	(17.35)	-0.030	(0.97)	-0.030	(0.97)
MKT	-0.432	(2.49)	0.068	(0.72)				
OPS & TECH	0.071	(0.56)	0.015	(0.34)	0.006	(1.53)	0.007	(1.94)
OR&MANSCI	0.318	(0.96)	0.075	(1.57)	0.058	(1.53)	0.048	(1.13)
ORG STUD	0.178	(1.42)	-0.110	(1.17)	0.021	(0.32)	0.034	(0.57)
PSYCH	-0.044	-(0.23)	0.164	(0.73)	0.227	(1.83)	0.227	(1.83)
PUB SEC	0.212	(1.06)	0.028	(0.53)				
SECTOR	0.049	(0.29)	-0.303	-(1.72)				
SOC SCI	0.129	(1.14)	0.002	(0.03)				
STRAT	0.335	(1.14)	0.082	(0.40)	0.452	(2.18)	0.452	(2.18)
TOUR-HOSP	0.263	(1.21)	-0.005	(1.08)				
No. significant (5%) gender coefficients		1		1		4		3
No. of sub-disciplines		24		24		14		14